



unidade **didáctica** 2

TRATAMENTO DOS DATOS ANALÍTICOS

M^a del Carmen Yebra Biurrun
Departamento de Química Analítica, Nutrición e Bromatoloxía
Facultade de Química



© Universidade de Santiago de Compostela, 2008

Deseño
Unidixital

Edita
Vicerreitoría de Cultura
da Universidade de Santiago de Compostela
Servizo de Publicacións
da Universidade de Santiago de Compostela

Imprime
Unidixital
Servizo de Edición Dixital da
Universidade de Santiago de Compostela

Dep. Legal: C 870-2008
ISBN 978-84-9750-936-7

ADVERTENCIA LEGAL: reservados todos os dereitos.
Queda prohibida a duplicación parcial ou total desta
obra, en calquera forma ou por calquera medio (elec-
trónico, mecánico, gravación, fotocopia ou outros) sen
consentimento expreso por escrito dos editores.

BIBLIOGRAFÍA BÁSICA

- ASSOCIATION OF OFFICIAL ANALYTICAL CHEMISTS (AOAC INTERNATIONAL). Official Methods of Analysis 18th Edition (Edition online) (2006): <http://eoma.aoac.org/>
- BASE DE DATOS DE MATERIAIS DE REFERENCIA CERTIFICADOS: <http://www.iaea.org/programmes/aqcs/>
- BIBLIOTECA DA UNIVERSIDADE DE SANTIAGO DE COMPOSTELA (USC): <http://www.busc.usc.es>
- HARRIS, D.C. (2006) Análisis Químico Cuantitativo, 3ª Ed., Reverté, Barcelona.
- INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY. Analytical Chemistry Division. http://www.ibiblio.org/pub/academic/chemistry/iupac_nm/V_compendium/
- MÉTODOS OFICIAIS DE ANÁLISE DE ALIMENTOS: <http://www.panreac.com/new/esp/publicaciones/publicaciones01.htm>
- MÉTODOS OFICIAIS DE ANÁLISE DE MOSTRAS MEDIOAMBIENTAIS: U.S. Environmental Protection Agency: <http://www.epa.gov/epaoswer/hazwaste/test/sw846.htm>
- MILLER, J.C., MILLER, J.N. (2002) Estadística y Quimiometría para Química Analítica, 4ª Ed., Prentice-Hall, México.
- OTTO, M. (1998) Chemometrics: Statistics and Computer Application in Analytical Chemistry, Wiley-VCH, Weinheim.
- RAMIS, G., GARCÍA, M.C. (2001) Quimiometría, Síntesis, Madrid.
- SKOOG, D.A., WEST, D.M., HOLLER, F.J., CROUCH, S.R. (2005) Fundamentos de Química Analítica, Thomson-Paraninfo, Madrid.

MATERIA: QUÍMICA ANALÍTICA II

TITULACIÓN: LICENCIATURA EN QUÍMICA

PROGRAMA XERAL DO CURSO

Localización da presente unidade didáctica

UNIDADE I. O PROCESO ANALÍTICO

ETAPAS FUNDAMENTAIS. RESOLUCIÓN DE PROBLEMAS ANALÍTICOS.

UNIDADE II. TRATAMENTO DOS DATOS ANALÍTICOS

TIPOS DE ERROS
EXACTITUDE E PRECISIÓN
ESTADÍSTICA APLICADA Ó ESTUDO DA EXACTITUDE E PRECISIÓN
GRÁFICOS DE CONTROL

UNIDADE III. A MOSTRA ANALÍTICA

MOSTRAXE E PROBLEMA ANALÍTICO. ERROS DE MOSTRAXE. PLAN DE MOSTRAXE. PREPARACIÓN DA MOSTRA PARA A ANÁLISE

UNIDADE IV. MÉTODOS ANALÍTICOS

CLASIFICACIÓN. CALIBRACIÓN. PARÁMETROS DE CALIDADE

UNIDADE V. MÉTODOS DE ANÁLISE GRAVIMÉTRICA

CLASIFICACIÓN. CARACTERÍSTICAS DAS REACCIÓN. FORMACIÓN E EVOLUCIÓN DOS PRECIPITADOS. CONTAMINACIÓN DOS PRECIPITADOS. PRECIPITACIÓN EN DISOLUCIÓN HOMOXÉNEA
TÉCNICAS E OPERACIÓNS DA ANÁLISE GRAVIMÉTRICA. CÁLCULOS GRAVIMÉTRICOS. APLICACIÓNS ANALÍTICAS

UNIDADE VI. INTRODUCCIÓN AOS MÉTODOS DE ANÁLISES VOLUMÉTRICAS

CARACTERÍSTICAS DAS REACCIÓNS UTILIZADAS EN VOLUMETRÍA
TIPOS DE VOLUMETRÍAS. MODOS DE VALORACIÓN. SUBSTANCIAS PATRÓN PRIMARIO. DISOLUCIÓNS PATRÓN. DETECCIÓN DO PUNTO FINAL CÁLCULOS VOLUMÉTRICOS

UNIDADE VII. VOLUMETRÍAS ÁCIDO-BASE

CURVAS DE VALORACIÓN. SISTEMAS MONOPRÓTICOS. SISTEMAS POLIPRÓTICOS. INDICADORES ÁCIDO-BASE. ASPECTOS PRÁCTICOS E APLICACIÓNS ANALÍTICAS

UNIDADE VIII. VOLUMETRÍAS DE FORMACIÓN DE COMPLEXOS

CURVAS DE VALORACIÓN. INDICADORES METALOCRÓMICOS
 ASPECTOS PRÁCTICOS E APLICACIÓNS: COMPLEXOMETRÍAS

UNIDADE IX. VOLUMETRÍAS DE PRECIPITACIÓN

CURVAS DE VALORACIÓN. SISTEMAS INDICADORES DO PUNTO FINAL. ASPECTOS PRÁCTICOS E APLICACIÓNS

UNIDADE X. VOLUMETRÍAS DE ÓXIDO-REDUCCIÓN

CURVAS DE VALORACIÓN. INDICADORES REDOX. REACTIVOS UTILIZADOS EN VOLUMETRÍAS REDOX. TRATAMENTO PREVIO DA MOSTRA CON OXIDANTES E REDUTORES. APLICACIÓNS

Valores para a variable F

Test de dúas colas. Nivel de significación (α): 5%. GL: graos de liberdade

		GL da varianza maior (numerador)														
GL da varianza menor (denominador)																
	1	2	3	4	5	6	7	8	9	10	12	15	20	60	∞	
1	647,8	799,5	864,2	899,6	921,8	937,1	948,2	956,7	963,3	968,6	976,7	984,9	993,1	1010,0	1018,0	
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,43	39,45	39,48	39,50	
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,34	14,25	14,17	13,99	13,90	
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,75	8,66	8,56	8,36	8,26	
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,52	6,43	6,33	6,12	6,02	
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,37	5,27	5,17	4,96	4,85	
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,67	4,57	4,47	4,25	4,14	
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,20	4,10	4,00	3,78	3,67	
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,87	3,77	3,67	3,45	3,33	
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,62	3,52	3,42	3,20	3,08	
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,28	3,18	3,07	2,85	2,72	
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,96	2,86	2,76	2,52	2,40	
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,68	2,57	2,46	2,22	2,09	
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27	2,17	2,06	1,94	1,67	1,48	
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05	1,94	1,83	1,71	1,39	1,00	

Valores para a variable F

Test de dúas colas. Nivel de significación (α): 1%. GL: graos de liberdade

GL da varianza menor (denominador)	GL da varianza maior (numerador)														
	1	2	3	4	5	6	7	8	9	10	12	15	20	60	∞
1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426	24630	24836	25253	25465
2	198,5	199,0	199,2	199,2	199,3	199,3	199,4	199,4	199,4	199,4	199,4	199,4	199,4	199,5	199,5
3	55,55	49,80	47,47	46,19	45,39	44,84	44,43	44,13	43,88	43,69	43,29	43,08	42,78	42,15	41,83
4	31,33	26,28	24,26	23,15	22,46	21,97	21,62	21,35	21,14	20,97	20,70	20,04	20,17	19,61	19,32
5	22,78	18,31	16,53	15,56	14,94	14,51	14,20	13,96	13,77	13,62	13,38	13,15	12,90	12,40	12,14
6	18,63	14,54	12,92	12,03	11,46	11,07	10,79	10,57	10,39	10,25	10,03	9,81	9,59	9,12	8,88
7	16,24	12,40	10,88	10,05	9,52	9,16	8,89	8,68	8,51	8,38	8,18	7,97	7,75	7,31	7,08
8	14,69	11,04	9,60	8,81	8,30	7,95	7,69	7,50	7,34	7,21	7,01	6,81	6,61	6,18	5,95
9	13,61	10,11	8,72	7,96	7,47	7,13	6,88	6,69	6,54	6,42	6,23	6,03	5,83	5,41	5,19
10	12,83	9,43	8,08	7,34	6,87	6,54	6,30	6,12	5,97	5,85	5,66	5,47	5,27	4,86	4,64
12	11,75	8,51	7,23	6,52	6,07	5,76	5,52	5,35	5,20	5,09	4,91	4,72	4,53	4,12	3,90
15	10,80	7,70	6,48	5,80	5,37	5,07	4,85	4,67	4,54	4,42	4,25	4,07	3,88	3,48	3,26
20	9,94	6,99	5,82	5,17	4,76	4,47	4,26	4,09	3,96	3,85	3,68	3,50	3,32	2,92	2,69
60	8,49	5,79	4,73	4,14	3,76	3,49	3,29	3,13	3,01	2,90	2,74	2,57	2,39	1,96	1,69
∞	7,88	5,30	4,28	3,72	3,35	3,09	2,90	2,74	2,62	2,52	2,36	2,19	2,00	1,53	1,00

ÍNDICE

Presentación	7
Os obxectivos	7
Os principios metodolóxicos	8
Os contidos básicos	9
1. O ERRO EN QUÍMICA ANALÍTICA.....	9
1.1. ERROS GROSEIROS OU ACCIDENTAIS	10
1.2. ERROS ALEATORIOS, INDETERMINADOS OU AO AZAR	10
1.3. ERROS SISTEMÁTICOS OU DETERMINADOS	10
1.4. DETECCIÓN E ELIMINACIÓN DE ERROS	11
2. PRECISIÓN	11
3. EXACTITUDE	12
4. ESTATÍSTICA E DATOS ANALÍTICOS	13
4.1. POBOACIÓN E MOSTRAS E PARÁMETROS ESTATÍSTICOS.....	13
4.1.1. Parámetros estatísticos que estiman o valor central.....	14
4.1.2. Parámetros estatísticos que estiman a dispersión dos resultados.....	14
4.2. DESCRIPCIÓN DUN CONXUNTO GRANDE DE DATOS: DISTRIBUCIÓN DE ERROS. FUNCIÓN DA PROBABILIDADE GAUSSIANA	15
4.3. COMPROBACIÓN DE HIPÓTESES ESTATÍSTICAS	18
4.4. TIPOS DE PROBLEMAS QUE REQUIREN A FORMULACIÓN DE HIPÓTESE E TESTS ESTATÍSTICOS PARA RESOLVELOS	19
4.4.1. Rexeitamento de valores sospeitosos (X_j).....	20
4.4.2. Comparación da precisión de dous conxuntos de datos. Comparación de varianzas	21
4.4.3. Comparación de medias.....	22
5. GRÁFICOS DE CONTROL.....	26
6. PRESENTACIÓN DOS DATOS ANALÍTICOS	29
Avaliación da unidade didáctica	30
Anexo 1: Exercicios numéricos	31
Anexo 2: Cuestións de autoavaliación	35
Anexo 3: Táboas de parámetros estatísticos	36
Bibliografía básica	42

Valores para a variable F

Test dunha cola. Nivel de significación (α): 5%. GL: graos de liberdade

		GL da varianza maior (numerador)														
GL da varianza menor (denominador)																
	1	2	3	4	5	6	7	8	9	10	12	15	20	60	∞	
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	246,0	248,0	252,2	254,3	
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,48	19,50	
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,57	8,53	
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,69	5,63	
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,43	4,36	
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,74	3,67	
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,30	3,23	
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,01	2,93	
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,79	2,71	
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,62	2,54	
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,38	2,30	
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,16	2,07	
20	4,35	3,49	3,10	2,87	2,71	2,60	2,49	2,45	2,39	2,35	2,28	2,20	2,12	1,95	1,84	
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,53	1,39	
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,32	1,00	

Valores para a variable F

Táboas de valores para a variable F. Test dunha cola. Nivel de significación (α): 1%. GL: graos de liberdade

GL da varianza menor (denominador)	GL da varianza maior (numerador)														
	1	2	3	4	5	6	7	8	9	10	12	15	20	60	∞
1	4052	5000	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6313	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,48	99,50
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,32	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,65	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,20	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,06	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	5,82	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,03	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,48	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,08	3,91
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,54	3,36
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,05	2,87
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,61	2,42
60	7,08	4,98	4,13	3,63	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	1,84	1,60
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,47	1,00

PRESENTACIÓN

Esta unidade didáctica é a primeira toma de contacto do alumnado co estudo do tratamento dos datos analíticos e a comprobación da súa calidade. Este estudo poderase completar máis adiante dentro do programa desta materia cando se presente o resto dos parámetros de calidade dos métodos analíticos. Posteriormente, no segundo ciclo da Licenciatura en Química, cando se curse a materia troncal Química Analítica Avanzada II, completárase o tema co estudo de ferramentas quimiométricas e co procedemento de validación dos métodos analíticos baseado nos parámetros de calidade. Tamén, na materia optativa Control de Calidade se estudaran en profundidade os gráficos de control de calidade e a súa utilidade práctica.

A importancia deste tema radica en que unha vez obtidos os datos analíticos, o primeiro aspecto que se ten que considerar é a verificación da calidade dos mesmos en termos de exactitude e precisión. A presenza ineludible de erros aleatorios impón o concepto de variabilidade dos datos, e obriga ó tratamento estatístico dos resultados co fin de estimar a fiabilidade e a robustez do método analítico. Ademais, a análise non se pode considerar finalizada ata que os datos obtidos non sexan procesados e expresados de maneira que poidan extraer toda a información que se require para a resolución do problema analítico.

Como a unidade didáctica anterior está dedicada a proporcionar unha visión global da Química Analítica centrada sobre o esquema do proceso analítico, a orde lóxica das seguintes unidades didácticas debería basearse na disposición cronolóxica das súas distintas etapas. Porén, esta unidade, que en realidade debería ser impartida ao final do programa, por ser o tratamento de datos unha das últimas etapas do proceso analítico, sitúase ao principio. A ruptura desta orde lóxica débese a un criterio práctico compartido pola xeneralidade dos docentes que imparten esta materia na nosa universidade e noutras, tanto españolas coma estranxeiras. Por exemplo, non é doado explicar a varianza asociada ó proceso de mostraxe (unidade didáctica III) se previamente non se introduce o concepto de erro, os tipos de erros e os parámetros estatísticos que se poden utilizar para avalialos. Do mesmo xeito, cando se define límite de detección e de cuantificación dun método analítico (unidade didáctica IV) será necesario explicar previamente os erros asociados á incerteza dunha determinación analítica.

OS OBXECTIVOS

- Identificar e analizar os distintos tipos de erros experimentais e a maneira de minimizalos e/ou eliminalos.
- Poñer de manifesto a magnitude do erro e a incerteza que pode estar implicada en cada unha das etapas da análise.
- Coñecer os parámetros de calidade asociados aos erros experimentais.
- Distinguir entre exactitude e precisión.

- Saber como avaliar no laboratorio a precisión e a exactitude dun método analítico.
- Localizar e seleccionar métodos oficiais de análise e materiais de referencia certificados.
- Manexar ferramentas estatísticas básicas necesarias para resolver distintos tipos de problemas relativos á exactitude e a precisión dos resultados e dos métodos analíticos.
- Comprender a necesidade de fornecer resultados analíticos de calidade: ser consciente da gran responsabilidade que implica dar un resultado dunha análise.

OS PRINCIPIOS METODOLÓXICOS

- A formulación metodolóxica basearase na participación do alumnado.
- Os principios teóricos e contidos fundamentais expóranse nas clases maxistras. Para iso, o profesor utilizará un ordenador para realizar presentacións.
- Propoñeremos actividades e exercicios prácticos que desenvolvan diferentes aspectos dos contidos que se teñan que aprender e desenvolvan as habilidades que se teñan que adquirir. Para iso, utilizaranse casos reais cos que se poderán tomar decisións fundamentadas e eficaces. Deste xeito, o alumnado terá a sensación de estar decidindo sobre cuestións reais similares ás que se enfrontará na súa futura vida profesional.
- As actividades desenvolvidas polo alumnado (de xeito individual ou nun grupo pequeno) para incidir sobre o cumprimento dos obxectivos do tema expóranse e discutíranse nas clases de seminario.
- Co obxecto de realizar cálculos estatísticos utilizaranse ferramentas informáticas. Na actualidade existen numerosos programas que permiten efectuar cálculos estatísticos (SPSS, Statgraphics, Minitab, etc.), aínda que a maioría deles presentan inconvenientes no seu uso, principalmente altos custos e complexidade no seu manexo. Fronte a isto, Microsoft Excel® aparece como unha das mellores alternativas, por canto dispoñemos inmediatamente del cando se instala a Suite Office. Adicionalmente, a facilidade no manexo de datos usando a folla de cálculo, fan de Microsoft Excel o software ideal para simplificar a aplicación de metodoloxías estatísticas.

Táboa de valores para a variable t

GL	NIVEL DE SIGNIFICACIÓN (α)					
	test de dúas colas			test dunha cola		
	10%	5%	1%	10%	5%	1%
1	6,31	12,71	63,66	3,08	6,31	31,82
2	2,92	4,30	9,92	1,89	2,92	6,97
3	2,35	3,18	5,84	1,64	2,35	4,54
4	2,13	2,78	4,60	1,53	2,13	3,75
5	2,02	2,57	4,03	1,48	2,02	3,36
6	1,94	2,45	3,71	1,44	1,94	3,14
7	1,89	2,36	3,50	1,42	1,89	3,00
8	1,86	2,31	3,36	1,40	1,86	2,90
9	1,83	2,26	3,25	1,38	1,83	2,82
10	1,81	2,23	3,17	1,37	1,81	2,76
11	1,80	2,20	3,11	1,36	1,80	2,72
12	1,78	2,18	3,06	1,36	1,78	2,68
13	1,77	2,16	3,01	1,35	1,77	2,65
14	1,76	2,15	2,98	1,35	1,76	2,62
15	1,75	2,13	2,95	1,34	1,75	2,60
16	1,75	2,12	2,92	1,34	1,75	2,58
17	1,74	2,11	2,90	1,33	1,74	2,57
18	1,73	2,10	2,88	1,33	1,73	2,55
19	1,73	2,09	2,86	1,33	1,73	2,54
20	1,72	2,08	2,85	1,32	1,72	2,53
30	1,70	2,04	2,75	1,31	1,70	2,46
60	1,67	2,00	2,66	1,30	1,67	2,39
120	1,66	1,98	2,62	1,29	1,66	2,36
∞	1,64	1,96	2,58	1,28	1,64	2,33

GL: graos de liberdade

ANEXO 3: TÁBOAS DE PARÁMETROS ESTADÍSTICOS

Táboa de valores para a variable Q

n	NIVEL DE SIGNIFICACIÓN (α)	
	0,05	0,01
3	0,970	0,994
4	0,829	0,926
5	0,710	0,821
6	0,628	0,740
7	0,569	0,680
8	0,608	0,717
9	0,564	0,672
10	0,530	0,635
11	0,502	0,605
12	0,479	0,579
13	0,611	0,697
14	0,586	0,670
15	0,565	0,647
16	0,546	0,627
17	0,529	0,610
18	0,514	0,594
19	0,501	0,580
20	0,589	0,567

Fonte: D.B. Rorabacher. Anal. Chem. 1991, 63, 139-146

OS CONTIDOS BÁSICOS

1. O ERRO EN QUÍMICA ANALÍTICA

A capacidade de proporcionar datos de calidade é de extrema importancia para os químicos analíticos en todos os eidos da análise. Deste xeito, as normativas nacionais e internacionais esixen aos laboratorios accións para asegurar a calidade da información que proporcionan. Estas medidas inclúen, entre outras, o uso de métodos de análise validados que dean lugar a resultados de calidade. Por iso, o tratamento e a avaliación dos datos analíticos inclúense como unha etapa máis do proceso analítico.

Esta etapa é a derradeira da estratexia da análise e consiste no procesado matemático dos datos para obter uns resultados que dean o valor máis probable da información buscada, así como a incerteza que o acompaña. Os contidos que se inclúen nesta unidade son os que se mostran no mapa de conceptos (Figura 1).

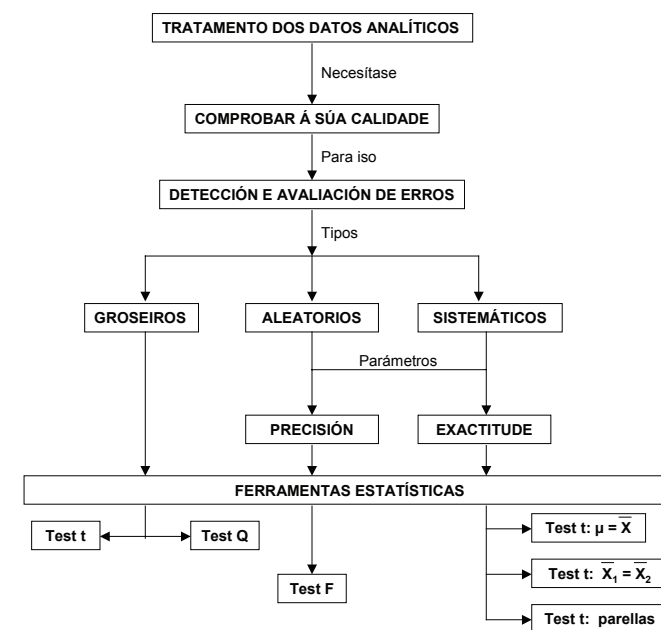


Figura 1. Mapa de conceptos dos contidos

Cando se realiza unha determinación analítica, o que se pretende é obter información veraz da propiedade que se está a medir (normalmente a concentración). Noutras palabras, que a concentración determinada do analito (valor experimental, X_i) sexa realmente a que existe na mostra (valor verdadeiro, μ). Deste xeito, o obxectivo de realizar unha determinación é obter unha estimación válida do valor verdadeiro. Con todo, este valor experimental pode levar asociado un erro cometido na determinación que faga que dito valor experimental non sexa o reflexo real do valor verdadeiro. Este erro (E)

será a diferenza en valor absoluto entre o valor experimental (X_i) e o valor verdadeiro (μ). Erro absoluto: $E_a = |X_i - \mu|$ e Erro relativo: $E_r = (E_a / \mu) 100$.

Xa que logo, durante a realización dunha análise cuantitativa é moi importante ter en conta todas as variables que poden afectar os resultados obtidos. Así, pódense interpretar para extraer as conclusións correctas que poidan resolver o problema que suscitou o comezo do proceso analítico. Os erros nas determinacións analíticas poden ser de tres tipos: groseiros, aleatorios e sistemáticos.

1.1. ERROS GROSEIROS OU ACCIDENTAIS

Son erros que se detectan con facilidade porque dan lugar a resultados moi diferentes de todos os demais datos dunha serie de medidas repetidas. Ocorren raras veces e poden causar que un resultado sexa moi alto ou moi baixo con respecto aos demais. Atribúense a erros persoais, descoidos na análise e manipulación da mostra e/ou reactivos, perdas de mostra, avaría total dun instrumento, caída ou vertido accidental dunha substancia noutra, etc.

1.2. ERROS ALEATORIOS, INDETERMINADOS OU AO AZAR

Dan lugar a que os datos ou medidas se dispersen máis ou menos simetricamente ao redor do valor medio ou central. Son debidos a flutuacións típicas de calquera proceso experimental de medida. Estes erros póñense de manifesto cando se realizan análises repetitivas sobre unha mesma mostra, obténdose resultados que oscilan ao redor dun valor central. As características deste tipo de erro son as seguintes:

- Non se poden predicir.
- Débense ao azar.
- Varían de signo e de magnitude (aínda que xeralmente esta é pouco importante).
- Seguen unha distribución Gaussiana.
- Obsérvase que cando aumenta o número de repeticións, uns erros compensan aos outros. Como resultado disto, o seu efecto sobre a estimación do valor central tende a diminuír.

Os erros aleatorios afectan á precisión dos resultados. É dicir, á concordancia entre os distintos valores obtidos ao realizar unha serie de medicións repetidas en condicións específicas.

1.3. ERROS SISTEMÁTICOS OU DETERMINADOS

Son debidos a alteracións operacionais ben definidas no proceso analítico. Provocan que a media dunha serie de datos sexa distinta do valor aceptado como verdadeiro (μ). Fan que todos os resultados sexan unidireccionais: erros sempre no mesmo sentido (positivos ou negativos). Os erros

ANEXO 2: CUESTIÓNS DE AUTOAVALIACIÓN

1. Diferenciar entre os tres tipos de erros experimentais.
2. Cal(es) dos seguintes apartados son correctos?
 - a. O erro experimental é sistemático ou aleatorio.
 - b. En ausencia de erro sistemático, a media está máis preto do valor verdadeiro que calquera medida individual.
 - c. Unha media de dúas determinacións é menos exacta que unha media de cinco determinacións.
3. Cando se analiza unha mostra por triplicado, o resultado obtido (\bar{X}) coincide co valor verdadeiro, quere dicir isto que non se cometeu ningún tipo de erro?
4. Se se repite a medición dun parámetro n veces do mesmo xeito. Se só se cometesen erros sistemáticos, como serían os valores obtidos en cada medición?
5. Pode ser unha desviación estándar igual a cero?
6. Se a S dun resultado é cero, que relación existe entre a media e as medidas realizadas?
7. Pode a media chegar a coincidir exactamente co valor verdadeiro ou aceptado?
8. É posible realizar unha determinación analítica sen cometer erros?
9. Especificar o tipo de erro que se comete nos seguintes procesos:
 - a. Precipitación incompleta nunha determinación gravimétrica.
 - b. Secado incompleto da mostra antes da pesada.
 - c. Erro na determinación do punto final dunha valoración.
 - d. Erro de indicador nun método volumétrico.
 - e. Perdas de precipitado nunha análise gravimétrica.
 - f. Calcinación incompleta dun precipitado antes da pesada.
10. Ao analizar unha mostra por triplicado, a S do resultado obtido (\bar{X}) é cero, quere dicir isto que non se cometeu ningún tipo de erro?
11. Que tipos de erros son os que se poden detectar utilizando a curva de Gauss?
12. Como afectan os erros aleatorios á curva de Gauss? E os sistemáticos?
13. Que erros son os causantes de que as medidas repetidas exactamente do mesmo xeito non sexan iguais?
14. Pode un resultado ser exacto e ao mesmo tempo pouco preciso?
15. Para determinar a precisión dun resultado basta con comparar os datos obtidos en análises repetidas pero, pódese determinar sempre a exactitude do resultado?
16. Cal(es) dos seguintes apartados son correctos?
 - a. Alta precisión implica alta exactitude.
 - b. A exactitude dunha medida experimental é 0,3%.
 - c. O peso medio dunha mostra é 7,1935 g.
 - d. A precisión dunha serie de datos é 0,7%.

10) Nun laboratorio de control de calidade do medio mariño analízase diariamente o chumbo nunha mostra control de mexillón. Os resultados obtidos foron os que se mostran na seguinte táboa.

tempo(días)	[Pb] µg/g	tempo(días)	[Pb] µg/g
1	0,15	11	0,16
2	0,17	12	0,13
3	0,14	13	0,15
4	0,15	14	0,16
5	0,13	15	0,15
6	0,15	16	0,17
7	0,17	17	0,12
8	0,13	18	0,11
9	0,13	19	0,10
10	0,18	20	0,09

Que se pode dicir da exactitude e da precisión do laboratorio de control de calidade do medio mariño en canto a determinación de chumbo en mexillóns?

Sol.: Antes do día 16 detéctase a presenza de erros aleatorios. A partir do día 16 confírmase a presenza de erros sistemáticos negativos.

sistemáticos afectan á exactitude dos resultados. É dicir, ao grao de concordancia entre o valor obtido experimentalmente e o valor verdadeiro.

Os erros sistemáticos poden ser erros do método, erros persoais e erros instrumentais, e os seus efectos dan lugar a:

- Erros constantes: teñen o mesmo valor absoluto independentemente da cantidade de analito ou tamaño da mostra.
- Erros proporcionais: aumentan ou diminúen dependendo da cantidade de analito ou tamaño de mostra utilizada na análise.

1.4. DETECCIÓN E ELIMINACIÓN DE ERROS

Os erros aleatorios non se poden evitar, só se poden avaliar estimando a precisión do procedemento analítico, e minimizar ao mellorar a precisión. Con todo, os erros sistemáticos pódense detectar avaliando a exactitude do procedemento analítico, e minimizar do seguinte xeito:

- Os erros persoais evítanse cun procedemento de traballo axeitado realizado por un analista con experiencia que traballe con coidado.
- Os erros instrumentais elimínanse realizando unha calibración e un mantemento periódico dos distintos compoñentes do instrumento analítico.
- Os erros do método evítanse cun desenvolvemento axeitado da metodoloxía analítica.

Actividades propostas

- Describir e dicir o tipo de erros que se poden cometer nunha determinación volumétrica.
- Describir e dicir o tipo de erros que se poden cometer nunha determinación gravimétrica.

2. PRECISIÓN

A precisión é unha medida do erro aleatorio asociado ao método analítico. Dise entón que un resultado é preciso se non presenta erros aleatorios ou polo menos se estes erros son aceptables. Expresa o grao de concordancia entre un grupo de resultados obtidos ao aplicar repetida e independentemente a mesma metodoloxía analítica a alícuotas da mesma mostra. Polo tanto, a precisión dun método analítico determínase no laboratorio realizando un número determinado de réplicas (normalmente once) da medida do analito nunha disolución patrón (disolución de concentración coñecida do analito preparada no laboratorio) ou nunha mostra.

Tamén é importante diferenciar entre reproducibilidade e repetibilidade, ambos conceptos relacionados coa precisión.

- Repetibilidade: Obtéñense resultados de ensaios independentes mediante o mesmo: método, laboratorio, analista, instrumento analítico, e nun intervalo curto de tempo. A repetibilidade consi-

dérase como unha medida da varianza interna e reflicte a precisión máxima que se pode obter cunha metodoloxía analítica.

- **Reproducibilidade:** Obtéñense resultados de ensaios independentes mediante o mesmo método aplicado á mostra en diferentes condicións, tales como: diferentes laboratorios, diferentes instrumentos analíticos, diferente analista, etc. Así, pódese establecer unha reproducibilidade entre días, analistas, laboratorios, etc.

Actividades propostas

- Detectar erros aleatorios na seguinte serie de datos experimentais resultado da medida da concentración de calcio en augas: 10,1; 10,0; 10,1; 10,2; 21,2; 10,0; 10,1 $\mu\text{g/mL}$. Comprobar como afectan estes erros á precisión dos resultados.
- Deseñar o xeito de abordar o estudo da precisión dun método proposto para a determinación de metais en pelo.
- Deseñar o xeito de abordar o estudo da precisión dun método proposto para a determinación de metais en auga de mar.

3. EXACTITUDE

A exactitude defínese como a proximidade na concordancia entre un resultado e o valor de referencia aceptado (valor verdadeiro, μ). Dise entón que un resultado é exacto se non presenta erros sistemáticos ou polo menos se estes son aceptables. A exactitude dun método analítico pódese avaliar no laboratorio de varias formas:

- Utilizando materiais de referencia certificados. Estes materiais teñen certificada a concentración dun ou varios compoñentes por un organismo competente e deben ter unha matriz similar á mostra a analizar. Xa que logo, cando se analizan, coñécese de antemán o valor verdadeiro (μ) do analito a determinar. Algún dos principais organismos que proporcionan materiais de referencia certificados son: Institute for Reference Materials and Measurements (IRMM) (Belgique), NIST: National Institute of Standards & Technology (USA), IAEA: International Atomic Energy Agency (Austria) e NRC: National Research Council (Canada).
- Utilizando métodos de referencia establecidos por distintos organismos oficiais. Estes métodos están o suficientemente probados para considerar que a concentración do analito obtida mediante a súa utilización corresponde co valor verdadeiro (μ).
- Realizando exercicios de intercomparación entre laboratorios. Neste tipo de exercicios, diversos laboratorios analizan un parámetro (ou unha serie de parámetros) nunha mostra e cada un proporciona un resultado. Normalmente, se a desviación estándar é aceptable, o valor medio obtido en todas as análises

É exacto o método espectrofotométrico para a determinación de ciclamato?

Sol.: O método espectrofotométrico é exacto ao 95% de probabilidade.

9) Nun laboratorio de control de calidade dunha industria láctea analízase diariamente o calcio nunha mostra control de leite. Os resultados obtidos foron os que se mostran na seguinte táboa.

tempo(días)	[Ca] $\mu\text{g/mL}$	tempo(días)	[Ca] $\mu\text{g/mL}$
1	0,5031	15	0,5042
2	0,5042	16	0,5040
3	0,5031	17	0,5035
4	0,5037	18	0,5030
5	0,5043	19	0,5037
6	0,5039	20	0,5037
7	0,5033	21	0,5027
8	0,5029	22	0,5024
9	0,5032	23	0,5022
10	0,5032	24	0,5018
11	0,5043	25	0,5014
12	0,5043	26	0,5000
13	0,5037	27	0,4910
14	0,5035		

Que se pode dicir da exactitude e da precisión do laboratorio da industria láctea?

Sol.: Antes do día 21 existen erros aleatorios. A partir do día 21 detéctase a presenza de erros sistemáticos negativos.

Datos Experimentais (n=6)							
[cafeína] µg/mL	15,5	14,9	15,0	15,2	14,8	15,0	15,5

O método novo é exacto?

Sol.: Método novo exacto, o valor verdadeiro atópase dentro do intervalo de confianza.

6) Un analista determinou quinina nun refresco utilizando o método oficial e utilizando outro método alternativo. Os resultados obtidos por ambos métodos foron os seguintes:

[quinina] µg/mL					
Método oficial	50,1	50,2	50,0	50,4	50,2
Método alternativo	50,6	50,3	50,4	50,2	50,3

É exacto o método alternativo para a determinación de quinina?

Sol.: O método alternativo para determinar quinina é exacto ao 95% de probabilidade.

7) Un analista determina a concentración de ácido benzoico nunha mostra por dous métodos distintos. Os resultados obtidos despois de facer seis réplicas da medida da mesma mostra son os que se mostran a continuación:

[ácido benzoico] µg/mL					
Método (1)	50,1	50,2	50,0	50,4	50,2
Método (2)	50,5	53,8	49,7	48,2	52,5

Son comparables as medias de ambos métodos?

Sol.: As medias destes métodos son comparables ao 95% de probabilidade.

8) Nun laboratorio preténdese avaliar un método para a determinación de ciclamato en bebidas baseado nunha técnica espectrofotométrica no ultravioleta. Para iso, analizáronse seis mostras distintas utilizando o método a avaliar. As mesmas mostras tamén se analizaron polo método oficial baseado nunha metodoloxía cromatográfica (HPLC). Os resultados obtidos expresados en µg/mL son os que se mostran na táboa seguinte:

Mostra	Método Oficial (HPLC)	Método Espectrofotométrico
refresco de laranxa	1,2	1,3
refresco de té	1,0	0,9
refresco de cola	3,7	3,9
auga tónica	5,5	5,1
refresco de limón	12,3	12,0
refresco de mazá	0,3	0,3

é o valor de referencia (valor verdadeiro, μ) para o parámetro determinado na mostra.

- Preparando unha mostra de referencia no laboratorio que conteña unha matriz o máis similar á mostra que se quere analizar e unha cantidade coñecida do analito que se quere determinar. Desta maneira, ao analizar o analito polo método analítico a avaliar coñécese o valor verdadeiro, μ .

Actividades propostas

- Detectar erros sistemáticos na seguinte serie de datos experimentais resultado da medida da concentración de calcio en augas: 10,1; 10,0; 10,1; 10,2; 21,2; 10,0; 10,1 µg/mL. Comprobar como afectan estes erros á exactitude dos resultados. $\mu = 10,1$ µg/mL.
- Localizar materiais de referencia certificados e métodos de referencia na rede.
- Deseñar o xeito de abordar o estudo da exactitude e precisión de métodos propostos para: a determinación de sacarina e ciclamato en edulcorantes, a determinación de metais en pelo, a determinación de metais en auga de mar, etc.

4. ESTADÍSTICA E DATOS ANALÍTICOS

A estatística permite sacar dos datos analíticos a máxima información e tomar decisións obxectivas que teñen unha alta probabilidade de ser correctas atendendo á súa calidade e uso.

4.1. POBOACIÓN E MOSTRAS E PARÁMETROS ESTADÍSTICOS

En estatística o termo "poboación" designa un grupo finito ou infinito de cousas con características comúns. En Química Analítica, sería o conxunto total de resultados que se poderían obter teoricamente se se realizara un gran número de análises.

A mostra é a parte observada dunha poboación, é dicir, un subconxunto desa poboación. En Química Analítica é un pequeno grupo de datos experimentais (obtidos no laboratorio) que se utiliza para extraer conclusións acerca da poboación da que proceden ou á que representan. Polo tanto, as mostras teñen que representar fielmente a poboación.

As leis da estatística aplícanse a unha serie de datos experimentais. Deste xeito, para usar estas leis en Química Analítica débese supoñer que a mostra é representativa da poboación de resultados. As estimacións estatísticas cunha mostra só poden ser feitas en termos de probabilidade. Os parámetros estatísticos representan as propiedades da mostra e serven para estimar as propiedades dunha poboación. Mediante a estatística obtense unha estimación de ata que punto os valores experimentais obtidos para as variables medidas representan os valores deses parámetros na poboación.

Na Táboa 1 móstrase a nomenclatura utilizada para os distintos parámetros estatísticos.

Parámetro	Poboación	Mostra	Relación
Número de obxectos	N	n	$n \leq N$
Media	μ	\bar{X}	\bar{X} é o estimado de μ
Desviación estándar	σ	S	S é o estimado de σ
Varianza	σ^2	S ²	S ² é o estimado de σ^2

Táboa 1. Nomenclatura utilizada para os distintos parámetros estatísticos

4.1.1. Parámetros estatísticos que estiman o valor central

Permiten coñecer a exactitude dos resultados. Os máis utilizados son os seguintes:

- **Media aritmética ou termo medio (\bar{X}):** serve para estimar o valor verdadeiro (μ) dunha poboación. X_i : cada un dos datos individuais; n: número de datos experimentais.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

- **Mediana:** valor central dunha serie de datos ordenados. Se o número de datos é par, é a media dos dous valores centrais.
- **Moda:** é o valor máis frecuente.

4.1.2. Parámetros estatísticos que estiman a dispersión dos resultados

Permiten coñecer a precisión dos resultados. Os máis utilizados son os seguintes:

- **Desviación estándar (S):** estima a dispersión dos resultados ao redor do valor medio. X_i : cada un dos datos individuais; \bar{X} : valor medio; n: número de datos experimentais; n-1: graos de liberdade (número de medidas independentes que se poden obter a partir dun conxunto específico de datos ou mostra).

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

ANEXO 1: EJERCICIOS NUMÉRICOS

1) Calcular a media, mediana e moda para os resultados experimentais da seguinte táboa:

	Datos Experimentais (n=9)								
[vitamina C] (µg/L)	15	16	15	17	19	17	18	16	17

Sol.: Media=17 µg/mL; mediana=17µg/mL; moda=17µg/mL

2) Calcular a desviación estándar, varianza, coeficiente de variación, desviación estándar da media e percorrido para os resultados experimentais da seguinte táboa:

	Datos Experimentais (n=9)								
[cloruro] (µg/mL)	15,1	15,3	15,8	15,7	15,3	15,6	15,7	15,1	15,4

Sol.: Media=15,4 µg/mL; S=0,265 µg/mL; varianza=0,07; CV=1,7%; S_m=0,09 µg/mL; P=0,7 µg/mL

3) Un analista efectúa 10 medidas nunha mesma mostra de auga para obter a concentración de cromo en auga de mar, e obtén os resultados que se mostran na seguinte táboa:

	Datos Experimentais (n=10)									
[Cr] (µg/mL)	10,1	11,0	10,3	10,5	10,5	10,4	10,5	10,4	10,5	10,4

Existe entre os resultados algún dubidoso?

Sol.: $X_q = 11,0$ µg/L: pódese rexeitar ao 95% de probabilidade.

4) Deséxase avaliar a precisión dun novo método analítico para a determinación de sacarina. Para iso analizouse unha mostra dun edulcorante por este método novo e polo método de referencia. Os resultados obtidos por ambos métodos móstranse na seguinte táboa:

	[sacarina] µg/mL					
Método de referencia	22,1	22,5	22,3	22,1	22,2	22,4
Método novo	22,1	22,7	22,3	22,4	22,0	

É preciso o método novo para a determinación de sacarina?

Sol.: Método novo preciso ao 95% de probabilidade.

5) Un químico desexa avaliar a exactitude dunha metodoloxía analítica nova utilizada para determinar cafeína en bebidas. Para iso, analizou utilizando esta metodoloxía un material de referencia, o cal certifica unha concentración de cafeína de 15,0 µg/mL. Os resultados obtidos foron os seguintes:

AVALIACIÓN DA UNIDADE DIDÁCTICA

Na avaliación do proceso de aprendizaxe do alumnado que se realizará ao longo de toda a Unidade Didáctica terase en conta:

- A observación do traballo diario na aula como unha información fundamental para a avaliación xeral dos alumnos.
- O traballo en grupo: participación e achega de cada membro, organización e calidade das actividades desenvolvidas.

Esta información completárase co traballo individual, onde se verá a superación das dificultades e os avances conseguidos. Para isto, cada alumno elaborará un cartafol no que incorporará: as actividades realizadas, os exercicios numéricos resoltos e a contestación ás cuestións de autoavaliación do tema.

O alumno será avaliado positivamente se:

1. Participa activamente nos seminarios onde se expoñen as actividades realizadas en grupo.
2. Realiza correctamente de xeito individual polo menos dúas das actividades propostas.
3. Resolve correctamente e/ou expón a maneira de solucionar o 75% dos exercicios propostos.
4. Contesta correctamente ao 75% das cuestións de autoavaliación.
5. Engade e realiza actividades e exercicios extras que lle axuden a mellorar a comprensión da Unidade Didáctica.

- **Varianza** (S^2): é o cadrado da desviación estándar (S).

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

- **Desviación estándar relativa ou coeficiente de variación** (CV): desviación estándar expresada como porcentaxe do valor medio.

$$CV = \frac{S}{x} \cdot 100$$

- **Desviación estándar da media ou erro estándar** (S_m): é o cociente entre a desviación estándar e a raíz cadrada do número de datos experimentais.

$$S_m = \frac{S}{\sqrt{n}}$$

- **Percorrido** (P): é a diferenza entre o valor máis alto ($X_{máx}$) e o máis baixo ($X_{mín}$) dunha serie de datos. $P = X_{máx} - X_{mín}$

Actividades propostas

- Exercicios numéricos 1 e 2 do anexo 1:
- Calcular parámetros estatísticos: cálculo manual.
- Calcular parámetros estatísticos: utilizando a folla de cálculo Microsoft Excel.

4.2. DESCRICIÓN DUN CONXUNTO GRANDE DE DATOS: DISTRIBUCIÓN DE ERROS. FUNCIÓN DA PROBABILIDADE GAUSSIANA

Na Táboa 2 móstrase unha serie de 50 datos obtidos na determinación de Zn nunha mostra dun mineral de cinc (expresados en %). Os datos dispérsanse debido á presenza de erros aleatorios.

Concentración de Zn nun mineral									
9,18	9,17	9,11	9,20	9,23	9,25	9,25	9,23	9,21	9,16
9,21	9,23	9,22	9,24	9,24	9,26	9,22	9,21	9,26	9,27
9,22	9,19	9,37	9,33	9,15	9,18	9,19	9,26	9,18	9,21
9,20	9,21	9,16	9,19	9,24	9,30	9,27	9,19	9,23	9,23
9,22	9,22	9,25	9,24	9,22	9,22	9,20	9,28	9,28	9,23

Táboa 2. Datos experimentais obtidos para a concentración de Zn nun mineral

Ao establecer unha táboa de frecuencias (Táboa 3), obtense o número de veces que se repite cada resultado (frecuencia absoluta, FA) e a porcentaxe con que se repite cada un deles (frecuencia relativa, FR).

% Zn	FA	FR (%)	% Zn	FA	FR (%)
9,11	1	2	9,23	6	12
9,15	1	2	9,24	4	8
9,16	2	4	9,25	3	6
9,17	1	2	9,26	3	6
9,18	3	6	9,27	2	4
9,19	4	8	9,28	2	4
9,20	3	6	9,30	1	2
9,21	5	10	9,33	1	2
9,22	7	14	9,37	1	2
Total: 50					

Táboa 3. Táboa de frecuencias dos datos da Táboa 2.

Cando se realiza a representación gráfica da distribución de frecuencias, obtense o histograma da Figura 2.

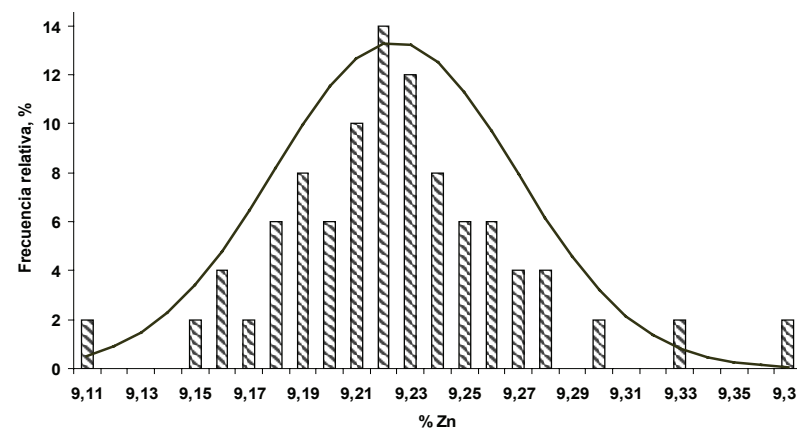


Figura 2. Histograma cos datos da Táboa 3.

Se se fai un número suficientemente grande de medidas, pódese esperar unha distribución de frecuencias (correspondente aos erros aleatorios) que se axustan a unha curva en forma de campá chamada Curva de Gauss ou curva normal de erro. A curva de Gauss pódese describir pola ecuación:

$$y = f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$$

ACTIVIDADES PROPOSTAS

- Exercicios numéricos 9 e 10 do anexo 1:
 - Avaliar a exactitude e a precisión dun método analítico usando cartas control.

6. PRESENTACIÓN DOS DATOS ANALÍTICOS

Cando se realiza a determinación dun analito nunha mostra, a súa análise como mínimo, débese realizar por duplicado. A incerteza destas medidas experimentais debe estar reflectida no resultado final. De maneira que á hora de expresar este resultado tense que proceder do seguinte xeito:

Resultado = (media \pm desviación estándar) unidades

Un aspecto relacionado coa presentación de resultados é o redondeo. O número de cifras significativas dadas indica a precisión do experimento. Na práctica fíxanse como cifras significativas todos os díxitos seguros máis o primeiro incerto. É importante destacar que os ceros á esquerda da coma non se consideran cifras significativas así como tampouco se consideran cifras significativas todos os ceros á esquerda do primeiro número distinto de cero. Este redondeo débese facer ao terminar os cálculos.

A continuación cítanse as regras do redondeo:

- Se o dígito seguinte á última cifra significativa é maior de 5: o número redondéase ao seguinte dígito maior.
- Cando é menor de 5: o número redondéase ao valor real da última cifra significativa.
- Se o último dígito é 5: o número redondéase ao seguinte dígito par máis próximo.

ACTIVIDADES PROPOSTAS

- Dada a seguinte serie de datos, obtidos na determinación de ácido cítrico en seis réplicas dun refresco de laranxa (n= 6): 55,3; 54,9; 21,1; 55,4; 55,0 y 55,1 $\mu\text{g/mL}$. Escribir a forma en que se debe dar o resultado da análise.

- Existirán erros sistemáticos se os dous resultados aos que se refire o apartado anterior están do mesmo lado da media.
- Existirán erros aleatorios se os resultados están nos distintos lados da media.
- O sistema estará fóra de control estatístico se un resultado está fóra do intervalo de acción. Neste caso, os resultados despois do último correcto débense rexeitar e débese actuar para corrixir o feito que causa o erro.

- Gráficos de Cusum:

Este tipo de gráficos consisten na representación da suma acumulada (C) das diferenzas entre cada valor (X_i) e o valor esperado en función do tempo ou do número de mostra (Figura 9).

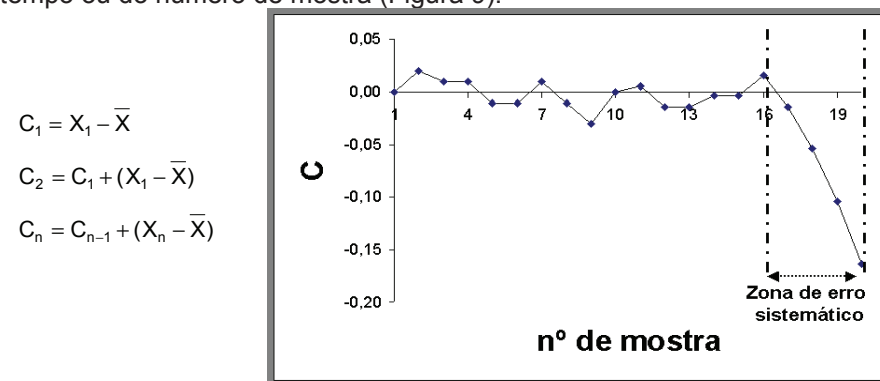


Figura 9. Gráfico de Cusum

Os gráficos de Cusum detectan con maior facilidade a existencia de erros sistemáticos, xa que este se vai sumar a cada seguinte valor de C, facendo que o valor de C se vaia afastando cada vez máis do punto cero.

Os erros aleatorios obsérvanse ao oscilar os valores de C ao redor do punto cero. Se $C=0$, significaría que non existen erros no proceso.

Onde μ é a media da poboación e σ é a desviación estándar. Para un conxunto limitado de datos (menos de 20 determinacións) no canto de poboación fálase de mostra: \bar{X} (media dos datos) e S (desviación estándar).

A forma máis ou menos esvelta das curvas de Gauss depende de σ (distancia desde μ ata o punto de inflexión da curva) (Figura 3).

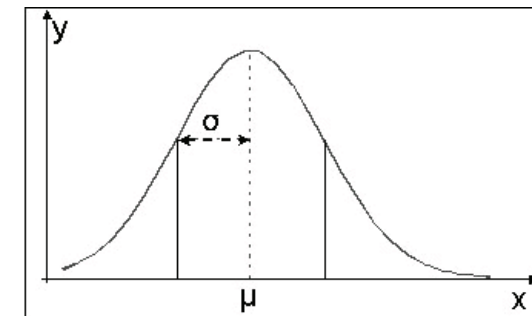


Figura 3. σ na curva de Gauss.

Nunha distribución de datos normal ou Gaussiana aproximadamente o 68% dos valores da poboación caen dentro do intervalo $\pm \sigma$ da media, preto do 95% dos valores sitúanse dentro de $\pm 2\sigma$ e máis do 99% dos valores atópanse dentro de $\pm 3\sigma$. Segundo sexa a forma da curva de Gauss, pódense detectar os tipos de erros da distribución normal. Así, na Figura 4 pódense observar tres curvas de Gauss coa mesma μ e distinta σ : curva máis ancha presenta maiores erros aleatorios (menor precisión), e polo contrario, a curva máis esvelta presenta menor erro aleatorio (maior precisión).

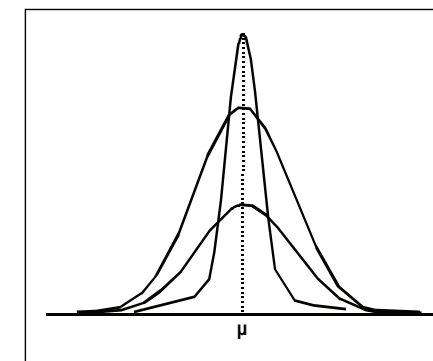


Figura 4. Curvas de Gauss co mesmo μ e distinto σ .

Na Figura 5 pódense ver tres curvas coa mesma σ , polo que as tres teñen os mesmos erros aleatorios (mesma precisión). Con todo, o valor de μ é diferente nos tres casos. Deste xeito reflíctense datos da medida da mesma propiedade na mesma poboación, isto tradúcese nunha falta de exactitude.

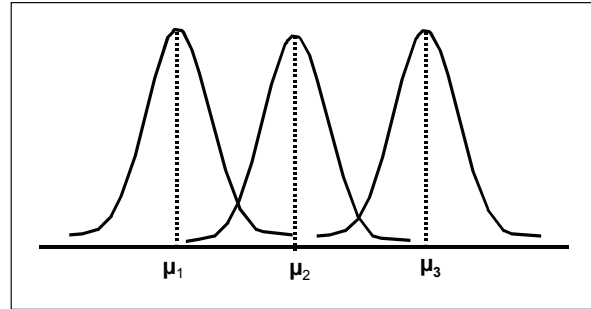


Figura 5. Curvas de Gauss co mesmo σ e distinto μ .

4.3. COMPROBACIÓN DE HIPÓTESES ESTADÍSTICAS

As medidas experimentais lévanse a cabo sobre mostras, non sobre poboacións. Xa que logo, existe un risco de tomar decisións erróneas se as características da mostra difiren das da poboación, débese cuantificar ese risco mediante a comprobación de hipóteses estatísticas. Estes ensaios estatísticos só son aplicables cando se supón que os datos teñen unha distribución normal ou Gaussiana.

A hipótese é a suposición que se realiza sobre un determinado suceso. Existen dous tipos de hipóteses:

- Hipótese nula (H_0): supón que os valores numéricos que se comparan son iguais. A diferenza atopada entre os valores experimental e teórico é debida á casualidade (só a erros aleatorios).
- Hipótese alternativa (H_1): é a hipótese contraria á hipótese nula. Establece que existe unha diferenza entre o valor experimental atopado e o valor coñecido para a poboación.

O nivel de significación (α) defínese como a probabilidade de rexeitar a hipótese nula cando en realidade é certa. Por conseguinte, o valor máis habitual para o nivel de significación é $\alpha = 0,05$: existen cinco ocasións de cada cen (5%) de rexeitar H_0 cando en realidade habería que aceptala.

Tamén, cómpre ter en conta que cando se aplica un test estatístico as hipóteses poden formularse de maneira que dean lugar a un test dunha cola ou a un test de dúas colas. O test de dúas colas cobre a posibilidade de que as diferenzas entre os valores que se comparan sexan positivas ou negativas. O risco de rexeitar H_0 repártese entre dúas colas de distribución normal da variable estatística. O test dunha cola cobre a posibilidade de que esas diferenzas teñan só un signo. O risco de rexeitar H_0 sitúase nunha cola da distribución normal.

Un modelo para tomar decisións estatísticas consta de varias etapas (Figura 6).

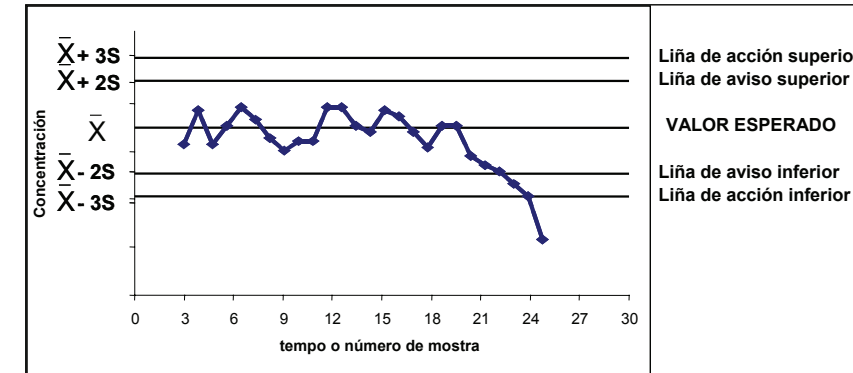


Figura 8. Gráfico de Shewhart

-Liña central: marca o valor esperado da variable medida.

-Límites de control: son os límites que se utilizan como criterio para tomar medidas ou para vulgar se un grupo de datos indica ou non unha falta de control. Os límites de control poden ser de dous tipos:

- Límites de aviso: son límites preventivos. Nunha distribución normal o intervalo entre a liña de aviso superior ($\bar{X} + 2S$) e inferior ($\bar{X} - 2S$) incorpora ao 95% dos resultados.

- Límites de acción: indican que hai que revisar o proceso. Nunha distribución normal o intervalo entre o límite de aviso superior ($\bar{X} + 3S$) e inferior ($\bar{X} - 3S$) incorpora ao 99,8% dos resultados.

O valor central e os límites débense fixar antes de iniciar o proceso de control de calidade. Daquela o valor central establécese mediante a análise previa dunha mostra control que presente as seguintes características: estabilidade, homoxeneidade, e con propiedades similares ás mostras reais que se queren controlar. O control establécese ao redor da media ou do valor verdadeiro. Así pois, a obtención de resultados fóra dos límites de control indica un empeoramento da precisión, ou que ata pode existir un erro sistemático.

Se o custo da análise non é moi elevado, pódense realizar gráficos nos que se representen as medias aritméticas de varias réplicas da mesma mostra. Neste caso, os límites de control son:

$$\text{Límite de aviso: } \bar{X} \pm \frac{3S}{\sqrt{n}} \quad \text{e límite de acción: } \bar{X} \pm \frac{2S}{\sqrt{n}}$$

Sendo n o número de réplicas realizadas para cada mostra.

As instrucións para interpretar os gráficos de control son as que se relatan a continuación:

- Non se fai ningunha acción se un resultado está dentro do intervalo de aviso e o seguinte se encontra fóra deste intervalo.
- Se dous resultados consecutivos están fóra do intervalo de aviso, pero dentro do de acción hai que revisar o proceso e actuar para evitar a tendencia.

Secuencia a seguir para realizar o test:

a. Calcular $t_{\text{experimental}} (t_{\text{EXP}})$: $t_{\text{EXP}} = \frac{\bar{X}_d - \sqrt{n}}{S_d}$

- b. Buscar o valor de t tabulado (t_{TAB}) para n-1 graos de liberdade, un nivel de significación α e test de dúas colas.
 c. Toma de decisión: Se $t_{\text{EXP}} < t_{\text{TAB}}$ acéptase H_0 . Non hai diferenzas entre os resultados dos dous grupos de datos. Se un dos métodos é un método oficial (método de referencia), o outro método é exacto.

ACTIVIDADES PROPOSTAS

- Exercicios numéricos 3-8 do anexo 1:
- Avaliar a exactitude e a precisión dun método analítico usando hipóteses estatísticas. Cálculo manual.
- Avaliar a exactitude e a precisión dun método analítico utilizando hipóteses estatísticas mediante a aplicación da folla de cálculo Microsoft Excel.

5. GRÁFICOS DE CONTROL

Os gráficos de control serven para controlar a calidade dos procesos analíticos en termos de exactitude e precisión. Estes gráficos detectan desviacións respecto ao valor real da medida efectuada \bar{X} (concentración de analito) ao longo do tempo, de maneira que é doado detectar calquera erro e realizar unha acción correctora o máis axiña posible.

Os gráficos de control máis utilizados son os Gráficos de Shewhart e os Gráficos de Cusum. A continuación descríbense cada un deles.

- Gráficos de Shewhart:

Consisten na representación da variable de interese (concentración) en función do tempo ou do número de mostra. Os elementos esenciais do diagrama (Figura 8) son os seguintes:

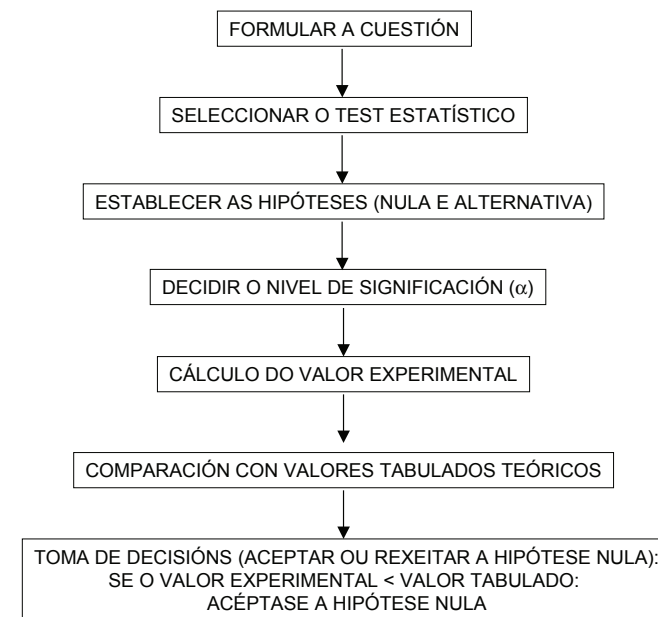


Figura 6. Sistemática a seguir na comprobación de hipóteses estatísticas.

4.4. TIPOS DE PROBLEMAS QUE REQUIREN A FORMULACIÓN DE HIPÓTESE E TESTS ESTADÍSTICOS PARA RESOLVELOS

A relación dos distintos test estatísticos aplicados ao tratamento dos datos analíticos obtidos experimentalmente móstranse na Figura 7.

A continuación descríbense cada un deles.

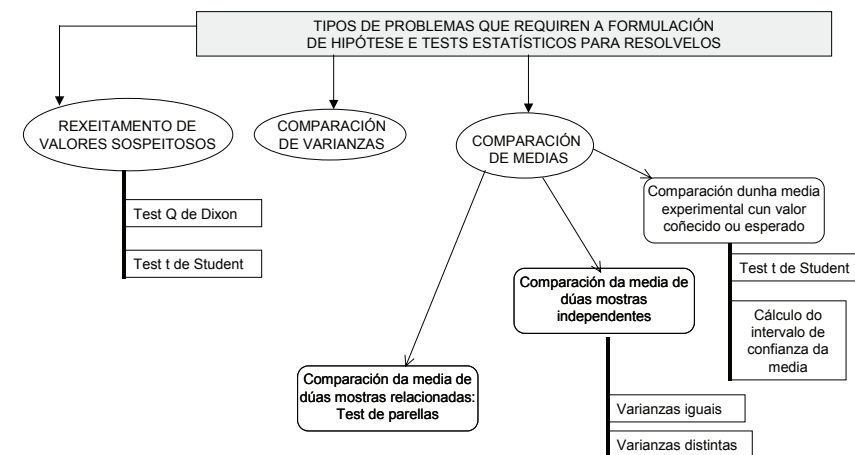


Figura 7. Tipos de problemas que requiren a formulación de hipótese e tests estatísticos para resolveolos.

4.4.1. Rexeitamento de valores sospeitosos (X_d)

Debidos a posibles erros groseiros. Existen dous test que se describen a continuación: test Q de Dixon e test t de Student.

- Test Q de Dixon:

É un test estatístico sinxelo e de uso xeneralizado cando se teñen n resultados de medidas experimentais e un dato sospeitoso X_d .

Hipóteses formuladas:

H_0 : X_d pertence á poboación

H_1 : X_d non pertence á poboación

Secuencia a seguir para realizar o test:

a. Ordenar de forma crecente ou decrecente os datos e localizar o valor dubidoso.

b. Calcular a diferenza en valor absoluto entre o valor dubidoso e o seu veciño máis próximo:

$$|X_d - X_{d-1}|$$

c. Calcular o percorrido P de toda a serie, é dicir, a diferenza entre o valor maior e o menor da serie de datos.

d. Calcular Q experimental (Q_{EXP}): $Q_{EXP} = \frac{|X_d - X_{d-1}|}{P}$

e. Buscar o valor de Q tabulado (Q_{TAB}) para n determinacións e un nivel de significación α .

f. Toma de decisión: se $Q_{EXP} < Q_{TAB}$ acéptase H_0 . O valor dubidoso pertence á poboación e non se debe rexeitar.

- Test t de Student:

Test estatístico baseado na distribución t de Student. Asímesa que entre n medidas hai un resultado dubidoso (X_d).

Hipóteses formuladas:

H_0 : X_d pertence á poboación de resultados con valor medio \bar{X} .

H_1 : X_d non pertence á poboación de resultados con valor medio \bar{X} .

Secuencia a seguir para realizar o test:

a. Ordenar de forma crecente ou decrecente os datos e localizar o valor dubidoso.

b. Calcular a media dos resultados (\bar{X}) e a desviación estándar (S) dos n-1 resultados (sen X_d).

Test de dúas colas: $H_0: \bar{X}_1 = \bar{X}_2$ e $H_1: \bar{X}_1 \neq \bar{X}_2$
Secuencia a seguir para realizar o test:

a. Calcular t experimental (t_{EXP}): $t_{EXP} = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$

b. Buscar o valor de t tabulado (t_{TAB}) para ν graos de liberdade, un nivel de significación α e test dunha ou dúas colas.

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_2^2/n_2)^2}{n_2+1}} - 2; \nu \text{ redondéase ao número enteiro máis próximo.}$$

c. Toma de decisión: se $t_{EXP} < t_{TAB}$ acéptase H_0 ; as medias son comparables.

4.4.3.3. Comparación da media de dúas mostras relacionadas (pares de valores, test t de parellas)

Realízase para comparar dúas metodoloxías analíticas, dous laboratorios ou dous analistas, etc. por medio do estudo de mostras que conteñen diferentes cantidades de analito. Así pois, obtéñense dous conxuntos de datos que se chaman *relacionados ou apareados*.

Este test estatístico aplícase co obxectivo de:

- Atopar se existen diferenzas significativas entre os resultados que se obteñen polos dous métodos, analistas, laboratorios, etc.
- Detectar erros sistemáticos asociados a unha concentración distinta.

A comparación efectuada baséase en calcular as diferenzas entre os pares de valores para comprobar se a media das diferenzas obtidas ao aplicar ambos métodos analistas, laboratorios, etc. é significativamente igual ou diferente de cero. Cando se utiliza para avaliar a exactitude, un dos métodos debe ser de exactitude contrastada (método oficial ou de referencia).

Hipóteses formuladas: test de dúas colas.

$H_0: \bar{X}_d = 0$: non hai diferenzas entre os resultados dos dous grupos de datos. Se un deles se obtivo utilizando o método de referencia o outro método é exacto.

$H_1: \bar{X}_d \neq 0$: hai diferenzas entre os resultados dos dous grupos de datos. Se un deles se obtivo utilizando o método de referencia o outro método non é exacto.

- **Caso 1:** se as varianzas son comparables.

Se aplicando o test F se acéptase a hipótese nula ($H_0: S_1^2=S_2^2$), entón acéptase que ámbas as dúas varianzas non son significativamente diferentes. Polo tanto, pódense combinar os datos de ámbas as dúas mostras nun só grupo para calcular unha varianza común para o novo conxunto de datos.

Daquela, calcúlase a desviación estándar ponderada (\bar{S}), e despois, a

$$\text{varianza conxunta } \bar{S}^2 : \bar{S} = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}; \bar{S}^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

Sendo n_1 e n_2 o número de réplicas realizadas por cada método e S_1 e S_2 , a desviación estándar do método 1 e do método 2, respectivamente.

Aplicación do test t de Student:

Hipóteses formuladas: test de dúas colas.

$H_0: \bar{X}_1 = \bar{X}_2$: medias comparables. Se unha delas se obtivo utilizando o método de referencia, o outro método é exacto.

$H_1: \bar{X}_1 \neq \bar{X}_2$: medias non comparables. Se unha delas se obtivo utilizando o método de referencia, o outro método non é exacto.

Secuencia a seguir para realizar o test:

a. Calcular t experimental (t_{EXP}):
$$t_{EXP} = \frac{\bar{X}_1 - \bar{X}_2}{\bar{S} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

b. Buscar o valor de t tabulado (t_{TAB}) para n_1+n_2-2 graos de liberdade, un nivel de significación α e test de dúas colas.

c. Toma de decisión: se $t_{EXP} < t_{TAB}$ acéptase H_0 : as medias son comparables. Se unha delas se obtivo utilizando o método de referencia, o outro método é exacto.

- **Caso 2:** se as varianzas non son comparables.

Se aplicando o test F se acepta a hipótese alternativa ($H_1: S_1^2 \neq S_2^2$), entón acéptase que ámbas as dúas varianzas son significativamente diferentes. Polo tanto, non se poden combinar os datos de ámbas as dúas mostras nun só grupo para calcular unha varianza común para o novo conxunto de datos. Logo, aplícase un test t aproximado cunha ou dúas colas.

Aplicación do test t de Student:

Hipóteses formuladas:

Test dunha cola: $H_0: \bar{X}_1 = \bar{X}_2$ e $H_1: \bar{X}_1 > \bar{X}_2$

c. Calcular $t_{\text{experimental}}$ (t_{EXP}):

$$t_{EXP} = \frac{\bar{X} - X_d}{S \sqrt{\frac{n}{n-1}}}$$

d. Buscar o valor de t tabulado (t_{TAB}) para $n-2$ graos de liberdade, un nivel de significación α e test de dúas colas.

e. Toma de decisión: se $t_{EXP} < t_{TAB}$ acéptase H_0 . O valor dudoso pertence á poboación e non se debe rexeitar.

4.4.2. Comparación da precisión de dous conxuntos de datos. Comparación de varianzas

Este test realízase cando se quere comparar a precisión entre dous métodos. Para iso débense realizar repeticións das análises con cada un destes métodos. Isto permite calcular as S de cada un deles e establecer así o seu grao de precisión (o máis preciso será o que presenta menor valor de S).

Cando se realizan as réplicas cos dous métodos analíticos, pódese calcular o seguinte: Método 1: n_1 réplicas-----> S_1 e S_1^2 , e Método 2: n_2 réplicas-----> S_2 e S_2^2

Así, para saber se existen diferenzas en canto a precisión entre ámbolos dous procedementos, procederase a comparar as varianzas (test F de comparación de varianzas).

Pódense dar dous casos:

- **Caso 1:** un dos métodos é de referencia (método oficial de análise), é dicir, de certificada precisión.

Hipóteses formuladas: test dunha cola porque non se espera obter unha incerteza menor que a obtida co método de referencia. S_r (desviación estándar do método de referencia) e S (desviación estándar doutro método que pode ser un novo método analítico).

$H_0: S^2 = S_r^2$: varianzas iguais. Como o método de referencia é preciso: método novo preciso.

$H_1: S^2 > S_r^2$: varianzas distintas: método novo non preciso.

Secuencia a seguir para realizar o test:

a. Calcular $F_{\text{experimental}}$ (F_{EXP}): $F_{EXP} = \frac{S^2}{S_r^2}$ sendo $F_{EXP} \geq 1$

b. Buscar o valor de F tabulado (F_{TAB}) para $n-1$ e n_r-1 graos de liberdade, un nivel de significación α e test dunha cola.

c. Toma de decisión: se $F_{EXP} < F_{TAB}$ acéptase H_0 . Varianzas iguais, como o método de referencia é preciso: método novo preciso.

- **Caso 2:** ningún dos métodos é de referencia.

Hipóteses formuladas: test de dúas colas. Faise para comprobar que dous métodos) difiren na precisión, ou na etapa previa no test de compara-

ción de medias. S_1 e S_2 (desviaciones estándares dos métodos 1 e 2, respectivamente).

$H_0: S_1^2 = S_2^2$: varianzas iguais. Os métodos presentan precisións similares.

$H_1: S_1^2 \neq S_2^2$: varianzas distintas. Os métodos presentan precisións diferentes.

Secuencia a seguir para realizar o test:

- Calcular $F_{\text{experimental}} (F_{\text{EXP}})$: $F_{\text{EXP}} = \frac{S_1^2}{S_2^2}$ sendo $F_{\text{EXP}} \geq 1$: no numerador pónese a varianza do método con maior desviación estándar (o menos preciso).
- Buscar o valor de F tabulado (F_{TAB}) para n_1-1 e n_2-1 graos de liberdade, un nivel de significación α e test de dúas colas.
- Toma de decisión: se $F_{\text{EXP}} < F_{\text{TAB}}$ acéptase H_0 . Varianzas iguais, métodos con precisións comparables.

4.4.3. Comparación de medias

Estes tests utilízanse para comprobar a exactitude dun grupo ou conxunto de resultados e poden ser de tres tipos:

- Comparación da media dunha mostra e a da súa poboación.
- Comparación de medias de dúas mostras.
- Comparación da media de dúas mostras relacionadas (pares de valores).

A continuación descríbense cada un deles.

4.4.3.1. Comparación da media dunha mostra e a da súa poboación

Este test estatístico aplícase cando se quere comparar unha media experimental cun valor coñecido (valor verdadeiro). Por conseguinte, serve para avaliar un método novo de análise. Esta avaliación pódese efectuar ao comparar o valor medio obtido experimentalmente cun método novo para a análise dun material de referencia certificado, coa media da poboación (valor certificado e coñecido dese material de referencia, μ).

Existen dous test que se describen a continuación: test t de Student e cálculo do intervalo de confianza da media.

- Test t de Student:**

Test estatístico baseado na distribución t de Student.

Hipóteses formuladas: test de dúas colas.

$H_0: \bar{X} = \mu$: media obtida polo método novo igual ao valor certificado. O método novo é exacto.

$H_1: \bar{X} \neq \mu$: media obtida polo método novo diferente ao valor certificado. O método novo non é exacto.

Secuencia a seguir para realizar o test:

- Calcular $t_{\text{experimental}} (t_{\text{EXP}})$: $t_{\text{EXP}} = \frac{|\bar{X} - \mu|}{S/\sqrt{n}}$
- Buscar o valor de t tabulado (t_{TAB}) para $n-1$ graos de liberdade, un nivel de significación α e test de dúas colas.
- Toma de decisión: se $t_{\text{EXP}} < t_{\text{TAB}}$ acéptase H_0 . O método novo é exacto.

- Cálculo do intervalo de confianza da media:**

O valor exacto da media dos datos dunha poboación (μ) nunca pode ser determinado exactamente porque esa determinación requiriría un número infinito de medidas. Con todo, a teoría estatística permite fixar límites ao redor dunha media determinada experimentalmente, de tal xeito que o valor verdadeiro estará entre eses límites cun determinado nivel de probabilidade. Estes límites chámanse límites de confianza e forman o chamado *intervalo de confianza*. O intervalo de confianza para a media dunha mostra vén dado por:

$$\mu = \bar{X} \pm t \frac{S}{\sqrt{n}}$$

Onde: \bar{X} = media das medidas da mostra, S = desviación estándar, n = número de determinacións ou tamaño da mostra, t = valor tabulado, para dúas colas da distribución t de Student, n-1 graos de liberdade e o nivel de significación α elixido.

4.4.3.2. Comparación da media de dúas mostras

Este test aplícase cando se quere comparar dous métodos de análise co obxecto de avaliar a exactitude dun método novo (e non se dispón de materiais de referencia certificados), para comparar como traballan dous laboratorios ou analistas distintos, etc. Cando se utiliza para avaliar a exactitude, un dos métodos debe ser de exactitude contrastada (método oficial ou de referencia).

Secuencia a seguir para realizar o test:

- Comparación das dúas dispersións ou varianzas: aplicar o test F de dúas colas, xa que interesa confirmar se as varianzas son iguais ou diferentes.
 - $H_0: S_1^2 = S_2^2$: varianzas comparables. Precisións iguais.
 - $H_1: S_1^2 \neq S_2^2$: varianzas no comparables. Precisións diferentes.
- Comparación das medias mediante un test t de Student. Pódense dar dous casos: